Control Groups in RCTs Supporting Approval of Drugs for Systemic Rheumatic Diseases, 2012-2022

Yang Liu, MD; Yan Xie, MD; Yunhe Qin, MD; Qibing Xie, MD; Xiaoyuan Chen, PhD

Abstract

IMPORTANCE Randomized clinical trials (RCTs) testing innovative drugs must strive to use optimal control groups to reflect the best available treatments. A comprehensive evaluation of the quality of control groups in pivotal RCTs supporting systemic rheumatic disease (SRD) drug approvals by the Food and Drug Administration (FDA) is lacking.

OBJECTIVE To examine the proportion of pivotal RCTs that used optimal control groups among RCTs supporting newly approved SRD drugs in the US over the past decade.

DESIGN, SETTING, AND PARTICIPANTS In this study, individual RCTs supporting SRD new drug approvals by the FDA between January 2012 and October 2022 were analyzed for design, study duration, control group, and primary end point. The quality of control groups was determined by comparison with published guidelines before and during the trial.

MAIN OUTCOMES AND MEASURES The primary measure was the proportion of RCTs using optimal control groups. Differences in response rate between investigating and control groups and the response rate of placebo control groups were also examined.

RESULTS Between January 2012 and October 2022, the FDA approved 44 SRD drugs, involving 65 pivotal RCTs. Overall, 16 RCTs used optimal control groups. In 55 trials, no active groups were used, and more than 80% of these trials were suboptimal (47 trials [85.5%]). Among 56 trials for systemic arthritis, 49 trials used suboptimal control groups, mainly placebo or dose-response controls (47 trials), with a few active controls (2 trials). Studies of other SRDs frequently used placebo or dose-response controls but were considered optimal controls (8 trials). There was significant improvement in response rates of investigating compared with placebo groups, with relative risk mostly exceeding 1.50 (range, 0.90; 95% CI, 0.69-1.17 for anifrolumab to 11.00; 95% CI, 2.69-44.96 for mepolizumab). In all placebo-controlled trials, the median (IQR) response rate in placebo groups was 26.0% (19.2%-32.3%).

CONCLUSIONS AND RELEVANCE These findings suggest that the quality of control groups in RCTs leading to SRD drug approval needs improvement and that despite challenges in translating scientific theories to clinical scenarios, it is crucial to consistently prioritize efforts to promote appropriate control group selection to ensure the accurate assessment of innovative drug efficacy.
Introduction

Systemic rheumatic diseases (SRDs), such as systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), psoriatic arthritis (PSA), and vasculitis, are heterogenous groups of diseases characterized by systemic inflammation and dysregulation of the autoimmune system. Interactions between genetic and environmental factors may contribute to triggering immune pathways, producing multiple autoantibodies and inflammatory factors and finally causing damage to multiple organs.1-3 Given the potential of severe and widespread organ damage, SRDs are associated with a considerable disease burden globally4-6 despite advances in the treatment landscape of SRDs.7-11 There is an urgent need for changes in current therapies and the development of new approaches.

Randomized clinical trials (RCTs) testing innovative drugs as the basis of regulatory approval and establishment of the standard of care should strive to use optimal control groups to reflect the best available treatment for patients with SRDs and provide appropriate treatment to patients designated to control groups. However, using questionable or suboptimal control groups is not uncommon in registration trials in various fields, such as oncology.12-14

This study aimed to systematically review the quality of control groups and the added clinical benefit of investing drugs in pivotal RCTs supporting the approval of SRD treatments by the US Food and Drug Administration (FDA). We also analyzed the potential justification and consequences from the perspectives of the regulatory agency and clinical practice.

Methods

This retrospective observational study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. Ethical approval and informed consent were deemed unnecessary as per the requirements of the Biomedical Ethics Committees in Beijing Tsinghua Changgung Hospital and West China Hospital because this is secondary research on existing public data that involves no human participants, individual information, or any specimens.

Search Strategy and Selection Criteria

We studied new drugs approved for treating SRDs in the US in the past decade. We defined the scope of SRDs as including RA, PSA, ankylosing spondylitis, nonradiographic axial spondyloarthritis, juvenile idiopathic arthritis, adult-onset Still disease (AOSD), SLE, systemic sclerosis, Sjögren syndrome, dermatomyositis or myositis, and systemic vasculitis. Approvals for treating single organs of specific SRDs, like those treating systemic sclerosis–associated interstitial lung disease, oral ulcers associated with Behçet disease, or lupus nephritis, were not included in this study.

We extracted all new drug applications and biologic license applications approved by the FDA for new molecular entities between January 1, 2012, and October 31, 2022, from the Pharmcube database (one of the most authoritative platforms of drug information), including the first approval and later new-efficacy indication or claim. The list was then verified using Center for Drug Evaluation and Research calendar year approval lists released on the FDA website. We excluded new drug application and biologic license application classifications pertaining to new formulations, new combinations, or other administrative or technical modifications to existing molecular entities that provided no additional efficacy evidence.

FDA clinical review documents and labels were collected from Drugs@FDA, the FDA-approved drugs database. Related literature, such as specific clinical trial publications, was also reviewed.

We extracted pivotal RCTs listed in FDA clinical review documents or from the label linked with specific approval if review documents were unavailable. Nonrandomized clinical pivotal trials were excluded from this study. For each product and pivotal trial identified, we evaluated details on the drug's basic information, design of the trial, trial population, duration of the trial, and results of primary efficacy end points.
Two authors (Y.L. and Y.X.) extracted and verified all data. Extracted data were tabulated using Excel software version 16.78 (Microsoft Corp).

**Assessing the Control Group**

Control groups in pivotal trials may include placebo control groups, different doses of the investigating drug (ie, dose-response control groups), and active control groups. We assessed the quality of control groups using 2 steps. First, 2 authors (Y.L. and Y.X.) performed a search of authoritative clinical guidelines, including American College of Rheumatology (ACR), European Alliance of Associations for Rheumatology, and other globally recognized guidelines, and published review articles for a particular disease to determine the standard of care therapy for a specific indication dating to before or during the trial study period. Reference clinical guidelines used to determine the standard of care are listed in the eAppendix in Supplement 1. Second, the 2 authors independently read published RCT data and protocols of pivotal trials to identify selected control groups and the specific study population and compared the control group selected with guideline recommendations to determine quality. We considered the quality of control groups to be suboptimal if restrictions were placed on the choice of control that excluded a recommended agent or if the control group but not the recommended agent was specified (eg, the control group was a single agent when doublet therapy was recommended). If there was no other recommended treatment option for the trial population, the placebo or dose-response control design should be optimal. Conflicts were resolved by referring to a third author (X.C.).

**Statistical Analysis**

Numerical data are presented with a median and an IQR. When presenting data on gained response rates of investigating drugs compared with control groups, we calculated the relative risk (RR) and 95% CI. For trials using flare occurrence rates as efficacy end points, we used the number of patients with no occurrence of disease to calculate the RR. Trials using time-to-event end points are not presented. For trials that involved multiple doses of investigating drugs, we calculated RRs using response rates of the approved dosage regimen. When more than 1 dosage regimen was approved, we used the pooled response rate. Statistical analyses and figure generation were performed using Excel software version 16.78, Prism software version 9.5.1 (GraphPad Software, LLC), and Stata statistical software version 15.0 (StataCorp). No formal comparisons were made in the analysis.

**Results**

**Characteristics of Drug Approvals and Pivotal Trials**

From January 2012 to October 2022, the FDA granted 44 approvals for treating SRDs, among which there were 7 initial approvals. Of these marketing authorizations, 3 approvals supported by nonrandomized pivotal trials and 4 not supported by added efficacy evidence were excluded.

We identified 65 pivotal RCTs leading to these approvals (eTable 1 in Supplement 1), including 17 trials for RA, 21 trials for PSA, 9 trials for ankylosing spondylitis, 4 trials for nonradiographic axial spondyloarthritis, 5 trials for juvenile idiopathic arthritis, 1 trial for AOSD, 5 trials for SLE, and 3 trials for systemic vasculitis. More than half of the trials lasted 1 to 2 years (36 trials [55.4%]), while 2 trials took more than 4 years to complete. Most trials enrolled 300 to 1000 patients (47 trials [72.3%]).

**Classification and Assessment of Control Group Used**

In 55 trials (84.6%), investigating drugs were tested against placebo or dose-response control plus placebo control groups without comparison with active treatments (eTable 2 in Supplement 1); 47 of these trials (85.5%) were suboptimal. Of all trials, 10 trials included active groups. However, most of these trials did not conduct statistical comparisons between investigating drugs and the active group. FDA approval pathways were all regular.
The distribution of control groups and their quality by disease types are depicted in Figure 1 and Figure 2. eTable 3 in Supplement 2 presents a list of pivotal trials with control group quality determination rationales. Overall, 16 RCTs used optimal control groups. Of 56 RCTs for systemic arthritis, including RA, PSA, ankylosing spondylitis, nonradiographic axial spondyloarthritis, and juvenile idiopathic arthritis, 49 trials used suboptimal control groups, including mostly placebo or dose-response controls (47 trials), as well as a few active and dose-response controls (2 trials). Studies of other SRDs, including AOSD, SLE, and vasculitis, commonly used placebo or dose-response controls but were all considered optimal controls (8 trials).

Figure 1. Pivotal Trials by Type of Control Group

Figure 2. Trials Using Suboptimal Control Groups by Disease Type

The type of control used in each trial is illustrated. Optimal control groups of rheumatoid arthritis (RA) and psoriatic arthritis (PSA) trials included those that involved optimal active control with concurrent questionable placebo. Optimal control groups of PSA and ankylosing spondylitis (AS) trials included those that involved optimal active control but compared the investigating drug only with concurrent placebo control. AOSD indicates adult-onset Still disease; AS, ankylosing spondylitis; JIA, juvenile idiopathic arthritis; nr-axSPA, nonradiographic axial spondyloarthritis; PSA, psoriatic arthritis; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; SRD, systemic rheumatic disease.
Among 40 trials testing biologics, 11 trials used optimal control groups, while among 25 trials testing targeted small molecules, 5 trials used optimal control groups. Among 3 trials that included patients newly diagnosed with SRD, all used optimal control groups.

Magnitude of Clinical Benefits Shown

Gained response rates of investigating drugs over treatments in control groups are shown in Figure 3 and the eFigure in Supplement 1. In general, we noticed significant improvements in response rates of investigating drug groups compared with placebo groups. The RRs of investigating drugs treating RA

The figure presents the relative risk (RR) on the response rate (primary efficacy outcome) across investigating groups compared with control groups in each pivotal trial, with error bars indicating 95% CIs. Sizes of boxes indicate the weight assigned to each individual trial in the final aggregated results. Drug-trial pairs are shown from the top down in the order of their approval sequence within the same indications. If unmarked, the pivotal trial used a placebo control. In trials that involved multiple doses of investigating drugs (dose-response control groups), RR was calculated using response rates of the approved dosage regimen. NA indicates not applicable.
each subtype of SRDs were within specific ranges, mostly greater than 1.50 (range, 0.90; 95% CI, 0.69-1.17 for anifrolumab to 1.10; 95% CI, 2.69-4.46 for mepolizumab). In 4 pivotal trials supporting investigating drugs for SLE, there were moderate increases in response rates. In 1 pivotal trial of anifrolumab, the response rate of the investigating drug was similar to that of the placebo group (RR, 0.90; 95% CI, 0.69-1.17), while in the other 2 trials, anifrolumab showed a response rate that was higher than the placebo group (RR, 1.95; 95% CI, 1.18-3.21 and RR, 1.53; 95% CI, 1.17-1.99).

Response Rates of the Placebo Control Group

Figure 4 plots response rates of placebo control groups in each pivotal trial by disease type. In all placebo-controlled trials, the median (IQR) response rate of the placebo groups was 26.0% (19.2%-32.3%). The median (IQR) response rate varied by disease type, with the lowest in vasculitis trials (8.5%; 3.0%-14.0%) and the highest rate in the AOSD trial (41.2%; 41.2%-41.2%).

Discussion

In this study, we summarized the quality of control groups in RCTs testing innovative drugs for SRDs. Our findings suggest that this quality needs to be improved in the future to facilitate novel drug marketing approval and clinical evidence guidance. Among these drugs, systemic arthritis products adopted more suboptimal control groups. We also calculated median increases in response rate compared with placebo or active groups and response rates of placebo groups to further investigate gained benefits associated with the innovative drugs tested. To our knowledge, this study is the first to thoroughly assess the quality of control groups and efficacy results in clinical trials that support the marketing of SRD drugs over a 10-year period, providing references for physicians and regulatory agencies.

The observation of suboptimal control groups used partly reflects obstacles to translating ideal scientific theories of testing the true efficacy of innovative drugs to the clinical scenario. Suboptimal controls are not limited to studies of SRDs. Studies indicate that control groups in anticancer drug trials are often of subpar quality. From an ethical standpoint, it is not acceptable to subject patients to placebos or therapies that are ineffective for a prolonged period, especially for diseases that warrant early control. It is important to avoid such practices at all costs. Regulatory agencies have made clear their stands in several guidelines, requiring that studies longer than 12 weeks should include an active comparator as the control or provisions for an escape to rescue treatment for patients with active disease. Add-on trials for patients having a flare and random withdrawal designs for patients achieving quiescence are also recommended. However, we found that many trials did not involve active treatment. Instead, a common approach was to permit early escape to rescue therapy after the primary efficacy follow-up. In such circumstances, patients were not treated

Figure 4. Response Rates of Patients Receiving Placebo

Trials in which the primary end point was not response rate are not shown here. AOSD indicates adult-onset Still disease; AS, ankylosing spondylitis; bars, median placebo response rates; dots, placebo response rates for individual trials; JIA, juvenile idiopathic arthritis; nr-axSPA, nonradiographic axial spondyloarthritis; PSA, psoriatic arthritis; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; SRD, systemic rheumatic disease; whiskers, IQRs.
as recommended in authoritative clinical guidelines. For example, patients with active psoriatic arthritis are recommended to switch to a tumor necrosis factor inhibitor, an interleukin-17 inhibitor (IL-17i), or an IL-12/23i biologic according to ACR 2018 guidelines published prior to the initiation of pivotal trials of risankizumab in 2019, in which placebo control was used with concomitant therapy with 2 or fewer conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) at protocol-approved doses. Moreover, it is difficult to interpret the advantage of investigating drugs over existing treatments. Indeed, our findings suggest that the approval of a single SRD drug was typically supported by several pivotal trials, and the magnitude of response rate improvement for patients receiving investigating drugs was prominent and relatively stable for a specific disease type, which reflects an overall favorable benefit-risk balance in regulatory decision-making.

The availability of treatment options has essential impacts on the strategy of choosing control groups. For systemic arthritis, such as RA, several existing drugs have been approved in the past years, and treatment paradigms are complicated. For such trials, it is important to select the appropriate control group based on the patient population being studied. Failure to do so could result in the use of suboptimal control groups. Active control groups can also be suboptimal. For example, in pivotal trials supporting upadacitinib to treat patients with active RA who may have had an insufficient response or intolerance to csDMARDs, methotrexate was the active control group. However, according to ACR guidelines, patients with insufficient response to 1 csDMARD should add another csDMARD or switch to biologics. For treating diseases like AOSD, SLE, and vasculitis, options are limited, which often leaves placebo or dose-response control as the only viable choices. Another challenge in selecting the best control group is the sparsity of high-quality evidence presented in reference guidelines.

In addition, we found that patients who received a placebo still showed positive responses in most trials. Interestingly, this is not the only study to have found this outcome. Barberio et al found moderate placebo responses in trials of drugs for irritable bowel syndrome with constipation or diarrhea, and placebo response rates calculated using different end points (pooled composite response, abdominal pain response, and stool response) were different. Even patients with cancer receiving placebo or no anticancer therapy were demonstrated to have a 1.95% overall response rate in a meta-analysis. There are several potential reasons to consider. First, some patients may have changed their background treatment regimen. For instance, in some trials, patients with RA who were previously intolerant or unresponsive were able to switch or add other csDMARDs, which could have added benefits to their treatment. This phenomenon also reminds investigators of the importance of choosing background therapy for an RCT to provide appropriate treatment for all participants and reduce the chance of suboptimal treatments. Second, patients included in RCTs were more likely to adhere to the regimen and were more frequently monitored than in real clinical practice. Third, prior treatments may have had some latent effects owing to disease characteristics. Most patients with RA who participated in the study were required to have used a previous treatment for at least 12 weeks, which is the recommended time for evaluating effectiveness. However, some patients are unable to achieve clinical remission or low disease activity within 3 months. Additionally, clinical guidelines recommend the treatment target as 6 months. All possible reasons highlight the importance of conducting well-controlled RCTs when testing drugs for SRDs. Furthermore, inclusion and exclusion criteria need to be carefully determined to enroll patients who are most likely to benefit from trials.

Using suboptimal control groups can have negative consequences. First, physicians may face difficulty in comparing newly approved drugs with existing treatments due to the lack of reliable evidence. Therefore, they may have to depend on biased indirect comparisons to determine the best treatment for their patients. This problem has lasted for more than a decade. In 2012, head-to-head trials of biologic DMARDs for RA and fair control groups were reported to be scarce. Additionally, drugs reviewed in our study were all granted regular approval, meaning there will be no required trials to test the effectiveness of investigating drugs relative to existing treatments. Second, network meta-analysis is a powerful tool to facilitate direct and indirect comparisons among drugs that
pertain to similar mechanisms of action. This becomes especially important when multiple new drug candidates are in similar stages of development, which frequently occurs when a new class of drugs is introduced to the market. Using optimal control groups across different clinical trials is expected to facilitate effective network meta-analyses, paving the way for future research. Third, from the patient perspective, receiving inadequate treatment during a clinical trial may result in a delay in receiving the proper treatment, potentially leading to deformities, even if rescue therapy or switching groups is an option. Many SRDs always cause irreversible damage to multiple organs, like bone erosion, lung fibrosis, and pulmonary hypertension, if treatment is delayed. In recent years, irrefutable evidence has emerged that supports using treat to target as an effective therapeutic strategy for patients. It is imperative that clinical remission is attained as swiftly as possible to enhance overall outcomes of the disease.

Globally, multiple regulatory and health policy communities have been actively engaging with patients throughout the entire drug development process. In the US, the FDA Patient-Focused Drug Development initiative was incorporated into the Prescription Drug User Fee Act V (reauthorized in 2012) and the 21st Century Cures Act (enacted into law in 2016). To better focus on patient needs in drug development, we recommend consistently prioritizing efforts toward promoting appropriate control groups for pivotal clinical trials. There have been heated discussions concerning this topic. For example, Kartolo et al raised the question of whether control groups of RCTs should even have an expiry date when a control treatment that is already proven to be inferior should be updated. When planning a pivotal RCT, it is important to use the best treatments available as the control group while carefully selecting inclusion and exclusion criteria. At least from the operational perspective, informed consent should reflect all the available choices for participants to review.

Limitations
This study has some limitations. First, the assessment of acceptable standard of care therapy can be subjective, although we tried to eliminate the subjectivity by individual review of the review package and clinical guidelines. Second, to minimize the impact of the availability of newly approved drugs when the trial was launched, we used clinical guidelines published before or during trials to determine the quality. Therefore, there are circumstances in which new drugs were approved for the trial population while the clinical guidelines were not yet updated. This method can lead to an inaccurate estimate of the control group quality. Third, there may be some conflicts between the determination of quality used in this study and additional regulatory considerations for drug approval, such as medical needs of the targeted population and the overall benefit-risk assessment of investigating drugs. Fourth, including all studies submitted to the FDA would provide a more comprehensive analysis. However, only the review package for approved drugs is available in the official database, which means that any trials related to drug marketing that were rejected cannot be accessed.

Conclusions
This study’s findings suggest that the quality of control groups in clinical trials leading to SRD drug approval needs improvement. Despite obstacles to translating ideal scientific theories of testing the true efficacy of innovative drugs to the clinical scenario, it is recommended to consistently prioritize efforts toward promoting appropriate control groups for pivotal clinical trials.
Open Access: This is an open access article distributed under the terms of the CC-BY License. © 2023 Liu Y et al. JAMA Network Open.

Corresponding Authors: Qibing Xie, MD, Department of Rheumatology and Immunology, West China Hospital of Sichuan University, No.37, Guoxue Alley, Chengdu 610042, China (xieqibing1971@163.com); Xiaoyuan Chen, PhD, Tsinghua Clinical Research Institute, School of Medicine, Tsinghua University, Beijing 100084, China (cyxy02648@mail.tsinghua.edu.cn).

Author Affiliations: Tsinghua Clinical Research Institute, School of Medicine, Tsinghua University, Beijing, China (Liu, Chen); Office of Clinical Trial Institute, Beijing Tsinghua Changgung Hospital, Beijing, China (Liu, Chen); Department of Rheumatology and Immunology, West China Hospital, Sichuan University, Chengdu, China (Y. Xie, Q. Xie); Pharmcube (Beijing) Co, Ltd, Beijing, China (Qin).

Author Contributions: Drs Liu and Y. Xie had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Drs Liu and Y. Xie contributed to the article equally.

Concept and design: Liu, Y. Xie, Q. Xie, Chen.

Acquisition, analysis, or interpretation of data: Liu, Y. Xie, Qin.

Drafting of the manuscript: Liu, Y. Xie.

Critical review of the manuscript for important intellectual content: All authors.

Statistical analysis: Liu, Y. Xie.

Administrative, technical, or material support: Y. Xie, Qin.

Supervision: Q. Xie, Chen.

Conflict of Interest Disclosures: None reported.

REFERENCES


